Synergy of Empirical Breeding, Marker-Assisted Selection, and Genomics to Increase Crop Yield Potential

Charles W. Stuber,* Mary Polacco, and M. Lynn Senior

ABSTRACT

This paper was presented as part of the symposium entitled "Post-Green Revolution Trends in Crop Yield Potential: Increasing, Stagnant or Greater Resistance to Stress." In this presentation, we have focused on (i) uses of marker technology in determining the genetic basis of phenotypic expression and the manipulation of phenotypic variation in plants. This included the use of markers in understanding heterosis, in attempts to improve hybrid predictions, in quantitative trait locus (QTL) identification and mapping, in marker-assisted selection (MAS), and in enhancing breeding success in the development of improved lines and hybrids; (ii) the role of genomics in developing a precise understanding of the genetic basis of phenotypic expression which will then provide more precision in the manipulation of phenotypic variation; and (iii) some attempts to integrate marker technology and genomics into empirical breeding strategies. In addition, we have focused on what has been successful as well as what has fallen short of expectations, and have suggested some of the possible reasons for the lack of success. Because of page limitations, we could not include an exhaustive review of the plant literature and have limited many of our examples to investigations in maize (*Zea mays* L).

The global ability to provide adequate amounts of food, feed, and fiber from domesticated crop plants has resulted largely from the collective empirical breeding efforts of farmers and plant breeders spanning many millennia. The continued increases in plant productivity have resulted from artificial selection, either conscious or unconscious, on the phenotypic expressions of the targeted species. Prior to the 20th century, plant breeding was largely an art with little or no knowledge of genetic principles. Although plant improvement since the rediscovery of Mendel's principles has involved both art and science, the contributions of science will undoubtedly assume a much greater role as new technology becomes more widely used and as additional gains in agricultural productivity are required to support a greater global population. New opportunities to use genotypic selection, or combinations of genotypic and phenotypic selection, to increase yield potentials are emerging at an ever-increasing pace.

On the basis of a comparison of 36 widely grown hybrids adapted to central Iowa and released at intervals from 1934 to 1991, Duvick (1997) reported that the increase in maize grain yield during that time span averaged nearly 74 kg ha^{-1} yr^{-1}. Hybrid comparisons were based on side-by-side trials, so all of the gain could be attributed to genetic improvement. Earlier studies indicated that genetic improvement usually accounted for about one-half of the total yield increase, with the remainder attributed to changes in cultural practices such as increased rates of mineral fertilizers and the use of herbicides for weed control and pesticides for control of insects and diseases. Duvick (1997) suggested that the increased grain yielding ability of these widely successful hybrids was due primarily to improved tolerance of abiotic and biotic stresses, coupled with the maintenance of the ability to maximize yield per plant under non-stress growing conditions.

Opportunities for gains resulting from changes in cultural practices are limited (particularly in the USA and other developed countries). Therefore, future gains in the productivity of most crops may depend almost entirely on genetic improvements. In fact, environmental concerns may cause a reduction in the use of agricultural chemicals and fertilizers. Also, many parts of the world may have limited supplies of such chemicals and plant nutrients. Therefore, plant breeders will need to develop and apply new technology (such as marker-assisted selection) at a faster pace to more effectively improve the yield potentials of crop plants for the ever increasing global human population as well as for the changes in consumer preferences.

Quantitative Traits and QTLs

A majority of economically important plant traits, such as grain or forage yield, can be classified as multigenic or quantitative. Even traits considered to be more simply inherited, such as disease resistance, may be "semi-quantitative" for which trait expression is governed by several genes (e.g., a major gene plus several modifiers). The challenge to use strategically new technology (such as DNA-based markers) to increase the contribution of "science" to the "art plus science" equation for plant improvement therefore applies to most, if not all, traits of importance in plant breeding programs. Although the focus of this symposium is on yield potential, that yield must be harvestable. Therefore, traits such as standability (or lodging resistance), disease resistance, and insect resistance must also be considered.

Historically, early researchers in quantitative genetics questioned whether the inheritance of these continuously distributed traits was Mendelian (Comstock, 1978). The answer to this question has major implications in the consideration of the use of markers for plant breeding programs. During the past century, both plant and animal geneticists have obtained convincing evidence that Mendelian principles apply to quantitative

C. Stuber, USDA-ARS and Dep. of Genetics, N.C. State Univ., Raleigh, NC 27695-7614; M. Polacco, USDA-ARS, Plant Genetics Res. Unit, Univ. of Missouri, Columbia, MO 65211; M. Senior, Novartis Agribusiness Biotechnology Research, Inc., 3054 Cornwallis Rd., Research Triangle Park, NC 27709. Received 28 Dec. 1998. *Corresponding author (cstuber@ncsu.edu).

Published in Crop Sci. 39:1571–1583 (1999).

Abbreviations: BAC, bacterial artificial chromosome; MAS, marker-assisted selection; QTL, quantitative trait locus, RFLP, restriction fragment length polymorphism; YAC yeast artificial chromosome.

as well as to qualitative traits. This evidence has also shaped the general model that embraces the multiple-factor hypothesis for quantitative traits (with genes located in chromosomes and hence sometimes linked, and incomplete heritability because of the contribution of environmental factors to total phenotypic variation).

If we agree that Mendelian principles apply to quantitative traits, we also need to define our concept of a QTL. Most geneticists and breeders consider QTLs to be chromosomal locations of individual genes or groups of genes that influence complex traits. Although it is often tacitly assumed that a QTL represents a single genetic determinant (or factor), there are examples of individual QTLs that have been resolved into multiple genetic factors by recombination (Graham et al., 1997; Yamamota et al., 1998). For the manipulation of the vast majority of QTLs in plant breeding programs, it may not be important to determine whether the QTL represents a single genetic factor or a cluster of tightly linked genes. However, if cloning of specific QTLs is paramount to their utilization, then the chromosomal location must be reduced to a manageable piece of DNA (Paterson, 1998b).

Some Opportunities and Challenges Facing Plant Breeders

The objective of this presentation is to focus on the synergy of empirical breeding, marker-assisted selection, and genomics to increase crop yield potential. Synergism has been defined as "united action—producing a greater effect than the sum of the various individual actions." We already have ample evidence for the success of empirical breeding in increasing crop productivity. Our goal then is to focus on opportunities to use marker-assisted selection, genomics, and other biotechnological innovations to enhance empirical plant breeding such that true synergism may be realized in the plant breeding arena.

Challenges facing plant breeders in their improvement programs are numerous, and many may be met with the development and application of new technology. Some of these challenges that have stimulated research in the application of marker technology and genomics relate to the following: (i) investigating and understanding the genetic and physiological bases of heterosis and prediction of hybrid performance, (ii) identification of useful genetic factors in divergent populations or lines (such as exotic accessions), (iii) introgression of desired genetic factors into breeding lines and breeding populations, (iv) enhancement of recurrent selection programs that are based on phenotypic responses, and (v) understanding and utilization of genotype $\times$ environment interaction. To meet these challenges, considerable emphasis has been placed on the development of new tools, such as DNA-based markers, with the major focus on the improvement of breeding precision and efficiency.

A multitude of investigations have been conducted on the inheritance of multigenic traits (using primarily classical biometrical methods), but plant breeders typically have little information on (i) the number of genetic factors (loci) influencing the expression of the traits, (ii) the chromosomal location of these loci, (iii) the relative size of the contribution of individual loci to trait expression, (iv) pleiotropic effects, (vi) epistatic interactions among genetic factors, and (vi) variation of expression of individual factors in different environments. The achievement of the maximum benefit from marker-based procedures for the manipulation and improvement of multigenic traits will require an increased understanding of the genetic and physiological bases underlying quantitative trait variation.

Marker-based technology already is providing scientists with a powerful approach for identifying and mapping QTLs and should ultimately lead to the development of a better understanding of genetic phenomena such as epistasis, pleiotropy, and heterosis. A number of recent investigations (particularly in maize; tomato, *Lycopersicum* spp.; and rice, *Oryza sativa* L.) are providing some clues to understanding such phenomena (e.g., Edwards et al., 1987; Stuber et al., 1987; Paterson et al., 1990, 1991; Abler et al., 1991; Koester, 1992; Edwards et al., 1992; Stuber et al., 1992; Eshed and Zamir, 1995; Li et al., 1997; Graham et al., 1997). It must be acknowledged, however, that studies such as these have identified and mapped only rather large chromosomal segments (in most cases probably 20–30 cM long). Although results from such studies may be adequate for many plant breeding endeavors, novel approaches will be necessary to identify individual genes and quantify individual gene action and interactions among genes.

Heterosis and Hybrid Predictions

Heterosis (or hybrid vigor) is a major reason for the success of the commercial maize industry as well as for the success of breeding efforts in many other crop and horticultural plants. Heterosis has been the topic of several conferences and symposia, the most recent being "The Genetics and Exploitation of Heterosis in Crops" held in Mexico City in August, 1997. Some progress has been made in understanding the genetic basis of heterosis. However, there is relatively little information regarding the biochemical, physiological, and molecular bases of this phenomenon.

The generation of inbred lines suitable for use in production of superior hybrids is very costly and requires many years in traditional empirical breeding programs. Much of the developmental effort is devoted to field testing of newly created lines in various single-cross combinations to identify those lines with superior combining ability.

Development of a reliable method for predicting hybrid performance without generating and testing hundreds or thousands of single-cross combinations has been the goal of numerous studies, using both marker data and combinations of marker and phenotypic data, particularly in maize. For example, in several earlier marker studies in maize, correlations between isozyme allelic diversity and grain yield were estimated in single-cross hybrids derived from commercially used lines (see

review by Stuber, 1999). Many of these studies used only 11 or fewer isozyme marker loci and 15 or fewer inbred parental lines. Consequently, the estimated correlations between isozyme allelic diversity and specific combining ability were low and nonsignificant in these studies. Even in a much larger study in which 100 maize hybrids derived from 37 elite lines were used to evaluate associations of hybrid yield performance with allelic diversity at 31 isozyme loci, an R^2 value of only 0.36 was reported by Smith and Smith (1989). Also, in another recent study, no association was found between hybrid grain yield and isozyme diversity in a study of six enzyme marker loci in 75 F_1 rice hybrids (Peng et al., 1988).

These retrospective correlation studies suggest that isozyme genotypes provide limited value in the prediction of hybrid performance in crops such as maize and rice. Several factors may contribute to this somewhat disappointing conclusion. For example, the low number of isozyme loci assayed in most of the studies would effectively mark only a small fraction of the genome. Therefore, only a limited proportion of the genetic factors contributing to the hybrid response would be sampled. More importantly, it is unlikely that these marker loci affect the phenotypic expression of the targeted quantitative trait directly; rather they serve to identify adjacent (linked) chromosomal segments. Allelic differences at marker loci *do not* assure allelic differences at linked QTLs. For a limited number of markers to be useful as predictors for hybrid performance, the effects of QTL "alleles" linked to specific marker alleles must be ascertained.

Also, it should be noted that the type of gene action associated with specific QTLs will affect the predictive value of linked marker loci. In maize populations developed from crosses of two inbred lines, it has been shown that the number of heterozygous marker loci is positively correlated with grain yield of F_2 plants or backcross families (Edwards et al., 1987; Stuber et al., 1992). These results corroborated other data that implicates dominant (or even overdominant) types of gene action as the predominate contributor to the expression of grain yield in maize. In such cases, marker allele diversity that reflects linked QTL allele diversity should be predictive of grain yield responses. However, for traits governed largely by additive gene action (this type of gene action might prevail for some loci affecting grain yield) the heterozygous QTL genotype would not be the most favorable. Again, as stated in the preceding paragraph, effective prediction of hybrid performance based on markers requires knowledge of QTLs linked to the markers.

In addition, it should be stressed that the level of linkage integrity must not be overlooked in the consideration of markers for hybrid predictions. For example, if the proposed hybrids are derived from lines produced from a randomly mated population, or if the lines comprise some subset of publicly available inbreds, then the associations between marker alleles and QTL alleles might be expected to be essentially random, i.e., near linkage equilibrium. For marker-based procedures to be successful for predictive or selective purposes for complexly inherited traits, such as grain yield, the genome should be well saturated with uniformly spaced markers and/or a high level of linkage disequilibrium must exist.

In the study discussed previously by Smith and Smith (1989), associations of grain yield with diversity of restriction fragment length polymorphism (RFLP) genotypes also were measured in the more than 100 hybrids derived from 37 elite maize inbred lines. Plots of F_1 grain yield against RFLP diversity, based on 230 marker loci, showed an R^2 value of 0.87. This value presents a striking contrast between the use of 230 RFLP marker loci versus 31 isozyme loci (with an R^2 value of 0.36) for the prediction of hybrid performance. However, it is important to note that even with 230 RFLP markers, yields varied from 8150 to 10 660 kg ha^{-1} (130–170 bushels/acre) for the subset of hybrids with the maximum detected "distance" (0.70–0.80 on a scale ranging from 0.00–0.80) between the parental lines. Most breeders would be working with similar subsets of largely unrelated lines for which, again, marker diversity alone does not appear to be very satisfactory for predictive purposes.

In another maize study, Melchinger et al. (1990) compared RFLP genotypes at 82 marker loci with field data on 67 hybrids reported earlier by Darrah and Hallauer (1972). Twenty inbred lines were involved in the parentage of the hybrids. They concluded that associations of hybrid yield, heterosis, and specific combining ability with multilocus heterozygosity of RFLP loci generally were too weak to be useful as a supplementary tool for predicting yield performance of crosses between unrelated lines. In addition, they concluded that for unrelated lines, genetic distance measures based on a large number of RFLPs uniformly distributed throughout the genome are not markedly superior to those based on a small number of isozymes for predicting hybrid yield. Thus, their results show that better marker coverage alone will not increase predictive power substantially. Melchinger et al. (1990) further state that "it seems necessary to employ specific markers for those segments that significantly affect the expression of heterosis for grain yield."

Dudley et al. (1992) reported a study for which the major objectives were to evaluate methods of using molecular marker data to: (i) identify parents useful for improving a single-cross hybrid and (ii) compare marker genotypic means measured at the inbred level to those measured at the hybrid level. Genotypic data from 14 isozyme and 52 RFLP marker loci were compared with field performance data from a diallel mating design of 14 maize inbreds in their investigations. They found that marker genotypic differences measured in inbreds were positively correlated with differences measured in hybrid backgrounds; however, these correlations were only slightly higher than those between phenotypic midparent and hybrid values. Their findings suggested that genotypic differences may be useful for preliminary selection of loci and alleles for possible improvement of hybrids but probably will not accurately predict final

performance of a hybrid. They concluded that number of unique alleles in a donor line was not a good measure for identifying lines that have value for improving a single cross and stated that "uniqueness of alleles does not necessarily indicate the presence of a favorable QTL." Thus, results from this study corroborated results from many earlier attempts to correlate marker allele diversity of parental lines with hybrid performance.

Bernardo (1994) evaluated the use of a best linear unbiased prediction of single cross performances based on (i) RFLP data on the parental inbreds and (ii) yield data on a related set of 54 single crosses. Sets of n predictor hybrids (where n = 10, 15, 20, 25, or 30) were chosen at random, and pooled correlations between predicted yields and observed yields of the remaining (54 − n) hybrids ranged from 0.65 to 0.80 (r^2 = 0.42–0.64). Although Bernardo concluded that single-cross yield can be predicted effectively based on parental RFLP data and yields of a related set of hybrids, these results are no better for hybrid predictions than those discussed above by Smith and Smith (1989).

The use of marker-aided prediction of advanced generation combining ability on the basis of data from early generation testcrossing was evaluated by Johnson and Mumm (1996). On the basis of a total marker score generated from regressing F_3 line testcross yields on the corresponding F_3 line RFLP genotypes, F_5 testcross yields were predicted with more accuracy than they would have been with only F_3 testcross yields as predictors. It was concluded that marker-aided prediction of advanced generation performance from early generation testcrosses was effective. These results were in agreement with those reported by Eathington et al. (1997).

QTL Identification and Mapping

QTL analysis has experienced a rapid evolution in the past decade which was made possible by numerous improved methods of molecular marker analysis. Advances in DNA technology in the past 20 yr, particularly restriction fragment hybridization analysis, polymerase chain reaction (PCR), and improved cytological analysis have revolutionized genetic analysis and generated new possibilities in the study of complex traits. As maps and markers progress from laboratory experimentation to tools that can be used by the practicing plant breeder, the technologies used for genotype analysis will become more robust and more reliable. Burow and Blake (1998) have provided a comprehensive compilation of the characteristics and merits of molecular tools now available and evolving for the study and manipulation of complex (quantitative) traits. However, because of the rapid advances in molecular technology, a compilation such as this will soon be outdated.

Mapping methods for identifying and locating QTLs are discussed very completely by Knapp (1994). Paterson (1998a) focuses on the technology of high resolution mapping which is expected to contribute substantially to such things as positional cloning of important genes and evaluating gene organization in divergent taxa as well as to applied objectives such as marker-assisted improvement of plants and animals. In addition, Kearsay and Farquhar (1998) have briefly reviewed methods currently available for QTL analysis in segregating populations and summarized some of the conclusions arising from such analyses in plant populations. They also have provided a summary of QTL properties from 176 trial-trait combinations in plants and details of the publications summarized can be found by accessing http://www.biology.bham.ac.uk/qtl-rev-papers/ (verified June 2, 1999).

As an example of the marker-facilitated research being conducted, we will focus on the maize genetics program at Raleigh, NC, in which QTLs have been identified and mapped in more than 20 populations (F_2, F_3, backcross, and recombinant inbred) derived from eight elite inbred lines and five inbred lines with a partial exotic (Latin American, expected to be 50%) component (Edwards et al., 1987; Stuber et al., 1987; Abler et al., 1991; Edwards et al., 1992; Stuber et al., 1992; Stuber, 1995; Koester et al., 1993; Ragot et al., 1995; Graham et al., 1997; and 1996, unpublished data). Both isozyme and RFLP marker loci were used in these studies, although the earlier studies used only isozymes. Measurements recorded on individual plants in the field experiments included dimensions, weights, and counts of numerous vegetative and reproductive plant parts as well as silking and pollen shedding dates.

Results from these studies showed that QTLs affecting most of the quantitative traits evaluated were generally distributed throughout the genome; however, certain chromosomal regions appeared to contribute greater effects than others to trait expression. In the earlier studies, not all of the chromosome regions were well marked, and presumably major factors also may have been segregating in regions of the genome devoid of marker loci in these studies.

Marker-Assisted Selection

In two of the early F_2 QTL detection studies conducted at Raleigh nearly 1900 plants were genotyped and evaluated for grain yield as well as for more than 80 quantitative traits in each population. Data from these F_2 populations were used to evaluate the efficacy of marker-facilitated manipulation of grain yield (Stuber and Edwards, 1986). Selections were based solely on the genotypes of the F_2 plants evaluated in the mapping studies; evaluations of selection response were then made on bulked progenies of these open-pollinated F_2 plants. In one of the selected populations, the mean of the increased-yield (based solely on genotypic values) entry was about 40% greater than the mean for the decreased-yield entry. Also, the mean yield of the increased-yield entry was about 20% greater than the mean of the unselected check (a sample of the open-pollinated population from the same F_2).

Field comparisons were made with phenotypic mass selection, and it was found that marker-facilitated selection (based on 15 isozyme marker loci which probably represent no more than 30–40% of the genome) was as

effective as phenotypic selection which would be expected to involve the entire genome. Furthermore, the results imply that a significant increase in the relative effectiveness of marker-based selection could be reasonably expected if the entire genome were marked with uniformly distributed loci (e.g., every 10–20 cM).

An investigation designed to develop improved inbred lines using QTL mapping information from two elite sweet corn breeding populations was reported by Edwards and Johnson (1994). They generated a selection index involving 34 traits and then correlated index performance with marker loci to determine which loci were associated with index performance. Significant gain in hybrid performance was evident in crosses of lines that had been generated by selection that was based only on marker genotypic information. They concluded that marker-facilitated selection allowed simultaneous gains for a number of traits, many of which require processing in a processing plant and are difficult and expensive to characterize.

Results from these population selection studies conclusively demonstrated that quantitative traits, such as grain yield, can be manipulated using only genotypic (marker) data. However, theoretical and analytical investigations by Lande and Thompson (1990) have shown that the maximum rate of improvement may be obtained by integrating both phenotypic and marker data. In their studies several selection indices were derived that maximize the rate of improvement in quantitative characters under several schemes of marker-assisted phenotypic selection (including the use of phenotypic data on relatives). They also analyzed statistical limitations on the efficiency of marker-assisted selection, which included the precision of the estimated associations between marker loci and QTLs as well as sampling errors in estimating weighting coefficients in the selection indices.

The Lande and Thompson (1990) investigations showed that (on a single trait) the potential selection efficiency by using a combination of molecular and phenotypic information (relative to standard methods of phenotypic selection) depends on the heritability of the trait, the proportion of additive genetic variance associated with the marker loci, and the selection scheme. The relative efficiency of MAS is greatest for characters with low heritability if a large fraction of the additive genetic variance is associated with the marker loci. Limitations that may affect the potential utility of marker-assisted selection in applied breeding programs include (i) the level of linkage disequilibria in the populations, which affects the number of marker loci needed, (ii) sample sizes needed to detect QTLs for traits with low heritability, and (iii) sampling errors in the estimation of relative weights in the selection indices.

The analytical approaches of Lande and Thompson (1990) focused on first generation selection. Succeeding studies have focused on the efficiency of MAS over several successive generations using computer simulations (Zhang and Smith, 1992, 1993; Gimelfarb and Lande, 1994a,b, 1995; Wittaker et al., 1995). Results from these studies showed that MAS could be more efficient than purely phenotypic selection in quite large populations and for traits with relatively low heritabilities. The simulations also showed that additional genetic gain provided by MAS, when compared with purely phenotypic selection, rapidly decreased when several successive cycles of selection had occurred, and that MAS may become less efficient than phenotypic selection in the long term. This situation becomes more acute when the effects associated with markers are not reevaluated at each generation.

Hospital et al. (1997) conducted computer simulations to study the efficiency of MAS based on an index combining the phenotypic value and the molecular score of each individual in the targeted population. In this case, the molecular score is computed from the effects attributed to markers by multiple regression of phenotype on marker genotype. Their results were consistent with earlier studies in that they also found that MAS may become less efficient than phenotypic selection in the long term. This is because the rate of fixation of unfavorable alleles at QTLs with small effects is higher under MAS than under phenotypic selection, and could be a consequence of the strong selection applied to QTLs with large effects under MAS in early generations. Hospital et al. (1997) pointed out, however, that this problem may be of little consequence in a practical breeding program because it takes place after a number of generations that is greater than the length of most breeding programs. They also indicated MAS "is of interest" when it is compared with purely phenotypic selection over several successive generations in a breeding program involving an alternation of generations with and without phenotypic selection, if heritability is high. In this situation, the effects attributed to markers are better estimated in the phenotypic evaluation step, so that selection on markers-only without phenotypic evaluation is then efficient in the next generation, even for small population sizes. In addition, the cost of MAS in this context is greatly reduced.

In a recent investigation, Knapp (1998) developed the theory for estimating the probability of selecting one or more superior genotypes by using MAS and included a parameter to estimate the cost efficiency of MAS relative to phenotypic selection. Depending on the level of the selection goal and the selection intensity, Knapp (1998) reported that a breeder using only phenotypic selection must test 1.0 to 16.7 times more progeny than a breeder using MAS to be assured of selecting one or more superior genotypes. Thus, MAS can substantially decrease the resources needed to accomplish a selection goal for a low to moderate heritability trait when the selection goal and selection intensity are high.

Although recent advances in molecular genetics have promised to revolutionize agricultural practices, Lande and Thompson (1990) state "There are, however, several reasons why molecular genetics can never replace traditional methods of agricultural improvement, but instead should be integrated to obtain the maximum improvement in the economic value of domesticated populations." Their analytical results, as well as the more recent computer simulations and the limited em-

pirical results, however, are encouraging and support the use of DNA-based markers to achieve substantial increases in the efficiency of artificial selection.

Other Examples of Uses of Markers to Enhance Breeding Success

Enhancement of Heterosis in an Elite Single Cross

The investigation reported by Abler et al. (1991) indicated that the maize inbred lines, Tx303 and Oh43, contained genetic factors that would be expected to enhance the heterotic response for grain yield in the B73 × Mo17 single-cross hybrid (Stuber and Sisco, 1991; Stuber, 1994, 1995). A further experiment provided appropriate data to identify six chromosomal segments in Tx303 and another six in Oh43 that (if transferred into B73 and Mo17, respectively) would be expected to enhance the B73 × Mo17 hybrid response for grain yield.

Three backcross generations (two marker-facilitated) were used for the transfer (introgression) of subsets of the identified chromosomal segments into the target lines, B73 and Mo17. This was followed by two generations of marker-facilitated selfing to fix the introgressed segments. The "enhanced" lines were then crossed in appropriate combinations and the "enhanced" single crosses were evaluated in replicated yield tests. On the basis of 4 yr of testing, yields of the best "enhanced" B73 × "enhanced" Mo17 hybrids exceeded the original B73 × Mo17 hybrid and high yielding commercial hybrids by 8 to 10% (628–1004 kg ha^{-1}).

Results from the transfer of the targeted segments from Tx303 into B73 and from Oh43 into Mo17 have demonstrated that marker-facilitated backcrossing can be successfully employed to manipulate and improve complex traits such as grain yield in maize. Not all of the six targeted segments have been successfully transferred into a single modified B73 or modified Mo17 line. There appears to be some indication that there may no advantage in transferring more than two to four segments. In fact, there is some indication that there could be a disadvantage. Increasing the number of transferred segments may be replacing the recipient genome with an excessive amount of linked donor chromosomal segments that could cause a deleterious effect. Also, epistatic interactions between a larger number of introgressed segments may result in a negative effect. In addition, favorable epistatic complexes in coupling phase (e.g., between recurrent parent alleles) could be disrupted. Further evaluations are necessary to determine the effects of larger numbers of transferred segments.

Marker-Facilitated Introgression (Backcrossing)

The preceding section outlined an investigation that demonstrated the use of marker-facilitated backcrossing; however, the procedures used were probably more complex than will be encountered in most plant breeding strategies. Several of the important traits that must be manipulated by plant breeders are more simply inherited than grain yield but still may involve the expression of several genes. For example, disease resistance and insect resistance frequently are controlled by only one or a few genetic factors. However, for many diseases and insect pests, resistance is considered to be a multigenic ("semi-quantitative") trait. For example, Bubeck et al. (1993) have shown that resistance to gray leaf spot (*Cercospora zeae-maydis* Theon & Daniels) in maize is based on at least four or five genes. Rector et al. (1998) reported one major QTL and two minor QTLs affecting corn earworm (*Helicoverpa zea* Boddie) resistance in soybean [*Glycine max* (L.) Merr.]. In addition, three major QTLs and two minor QTLs affecting antixenosis resistance to corn earworm were found in soybean by Rector et al. (1999).

In most cases, the first steps in a marker-based introgression program are the identification and mapping of the genes (more realistically, chromosomal segments) targeted for transfer to the desired line or strain. As suggested by Kearsey and Farquhar (1998), breeders may not need to know the locations of their targeted QTLs with very great accuracy. They will mainly be interested in incorporating (into elite lines) those QTLs which have a large effect, and which may have been missed by conventional selection procedures. Once the appropriate analyses have been performed to identify the genes of interest in the resource (perhaps, exotic) strain, as well as linkages to resource-specific marker alleles, repeated backcrossing to the recipient line or cultivar—choosing in each cycle only those backcross progeny with desired linked marker alleles—will provide effective introgression of the desired genes of interest into the recipient line. As was demonstrated in the previous section, marker-assisted selection against unwanted chromosomal regions from the donor (reducing linkage drag) will expedite the introgression process.

Beckmann and Soller (1986) calculated the frequency of a favorable allele following one to six backcross generations, with and without selection for a linked marker allele, including selection for a pair of bracketing marker alleles. They found that the frequency of the introgressed favorable allele after three generations of backcrossing is 0.66 for the single marker and 0.85 for bracketing markers (with recombination of 0.40 between markers). These frequencies are in striking contrast to the 0.06 with no markers after three backcrosses and only 0.01 after six backcrosses. Also, with marker-assisted introgression, frequencies for the introgressed alleles are sufficiently high that two, three, or even more alleles could be readily introduced and brought to fixation in a given breeding cycle. As stated by Beckmann and Soller (1986), "Without marker assistance, a great many backcross products will have to be screened for the introduced trait, even in the case of one introduced allele, due to the extreme rarity of backcross products carrying desired exotic alleles."

Near-isogenic Lines (NILs) as a Breeding Tool

The results discussed earlier showed that the enhancement of lines B73 and Mo17 was successful, but the procedure for the development of the enhanced lines

(NILs) was very inefficient and would not be recommended for a practical breeding program. That procedure depended on the identification of the targeted segments (containing the putative QTLs) prior to transfer to the recipient lines. In our maize research program at Raleigh, we have outlined and tested a marker-based breeding scheme for systematically generating superior lines without any prior identification of QTLs in the donor source(s). The identification and mapping of QTLs in the donor is a bonus obtained when the derived NILs are evaluated. Choice of the donor usually will be based on prior knowledge of its likely potential for providing superior genetic factors, and, in maize, may involve appropriate heterotic relationships.

The procedure involves generation of a series of NILs by sequentially replacing segments of an elite line (the recipient genome that is targeted for improvement) with corresponding segments from the donor genome. The objective is to generate a set of NILs containing, collectively, the complete genome of the donor source, with each NIL containing a different chromosomal segment from the donor. Marker-facilitated backcrossing, followed by marker-facilitated selfing to fix introgressed segments, is used to monitor the transfer of the targeted segments from the donor and to recover the recipient genotype in the remainder of the genome. The number of backcrosses required will depend on the number of evaluations that can be made in the marker laboratory. As few as two backcross generations and one selfing generation will suffice if the laboratory resources are adequate to handle the required number of plant samples.

In maize, the NILs are then crossed to an appropriate tester(s) to create hybrid testcross progeny that are evaluated in replicated field trials (with appropriate checks) for the desired traits. (The NILs would be tested per se for crops such as soybean and wheat, *Triticum aestivum* L.) The superior performing testcrosses will be presumed to have received donor segments that contain favorable QTLs. Thus, QTLs are mapped by function, which should be an excellent criterion for QTL detection. The breeding scheme not only creates enhanced elite lines that are essentially identical to the original line, but it also provides for the identification and mapping of QTLs as a fringe benefit with no additional cost. Obviously, the scheme is based on having a reasonably good marker map with distinct alternate alleles in the donor and recipient lines.

This breeding strategy should be an excellent procedure for tapping into the potential of exotic germplasm. Furbeck (1993) used this procedure to develop a set of 149 NILs using the elite line, Mo17, as the recipient line and B73 as the tester line. An exotic population, derived from the Brazilian racial collection Cristal (MGIII) and the Peruvian collection Arizona (AYA41), was used as the donor. Figure 1 shows some of the significant introgressed segments and traits involved. By this procedure, positive segments (such as the segment associated with *Phi1* on chromosome 1) from the exotic line are immediately available in an adapted background (Mo17) for further breeding use. Also, and equally important, segments with negative effects—such as those associated with *Dia1* and *Dia2*, both -596 kg ha^{-1} (-9.5 bushels/acre)—are eliminated. Moreover, because each segment was incorporated independently, the detection of positive effects is not biased by adjacent negative segments—e.g., *Phi1*, $+659$ kg ha^{-1} ($+10.5$ bushels/acre), and *Dia2*, -596 kg ha^{-1} (-9.5 bushels/acre), on

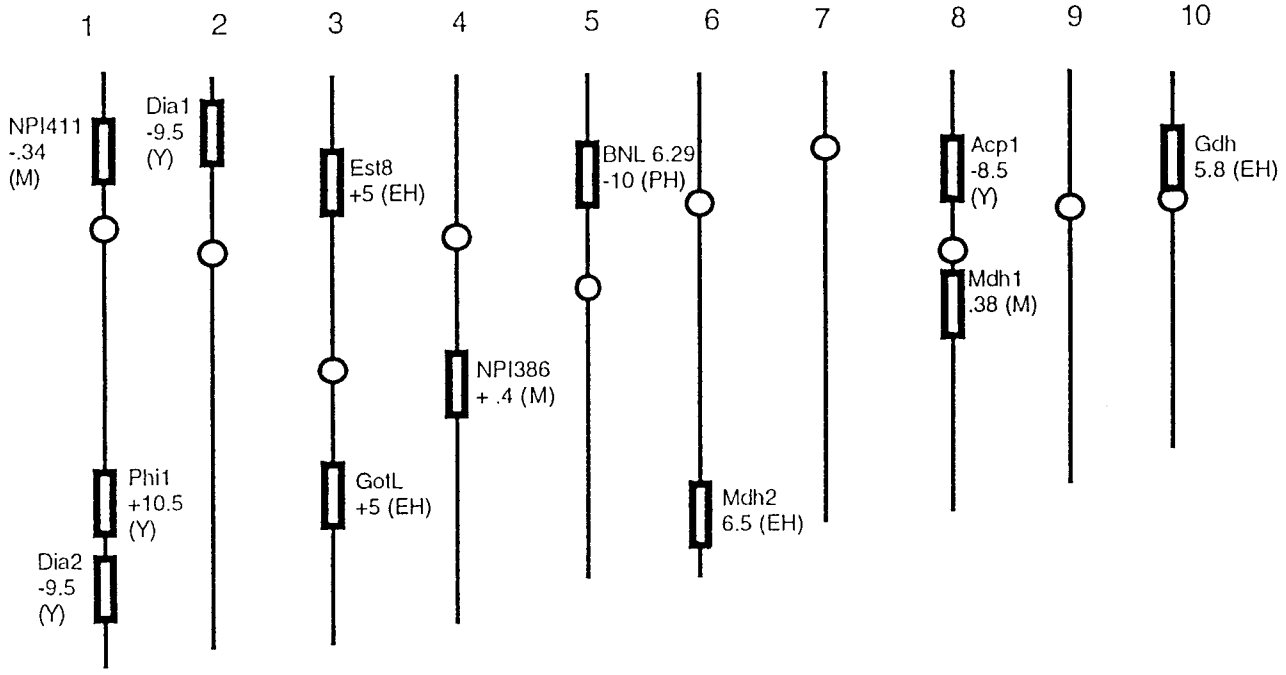

Fig. 1. Maize chromosome map showing locations of positive and negative effects on several traits derived from segments transferred from an exotic population (Cristal × Arizona) into the elite inbred line Mo17 with B73 used as the tester line. Y = yield (bushels/acre); M = moisture(%); EH = ear height (cm); PH = plant height (cm).

chromosome 1. These adjacent segments would likely segregate together in a traditional breeding program and would effectively cancel each other.

Brown et al. (1989) used a somewhat similar marker-facilitated backcrossing scheme to transfer isozyme-marked segments from wild barley (*Hordeum spontaneum* K. Koch) into an elite barley (*H. vulgare* L.) cultivar. Each of the 84 NILs was then made homozygous for a single isozyme-marked segment with two selfing generations. After evaluating these lines, per se, in the field, they also concluded that this was a useful approach for identifying QTLs for improving yield in divergent germplasm. Eshed and Zamir (1995) have also very effectively used the NIL breeding procedure to extract favorable genetic factors from a wild species of tomato (*Lycopersicum* spp.).

A major advantage of this NIL approach is that once a favorable QTL has been identified, it is already fixed in the elite recipient line and the breeding work is essentially completed. Also, because only a small segment of the genome of the recipient line has been modified, the enhanced line is nearly identical to the original line and the amount of field testing required is minimal. In addition, lines with favorable QTL alleles can be easily maintained and then used for pyramiding several favorable QTL alleles into a single line. A possible disadvantage of this approach is that favorable epistatic complexes between QTLs may not be identified. However, there is little experimental evidence documenting the occurrence of such epistatic interactions.

Genomics

Once a QTL has been defined, and closely linked markers have been identified, the region can be moved among lines by standard marker-facilitated protocols. Precise manipulation of the phenotype, however, requires identification of the gene(s) controlling the trait of interest. Several methods can be used to approach this problem. One method involves mapping a series of large, overlapping genomic clones, such as bacterial artificial chromosomes (BACs) or yeast artificial chromosomes (YACs), to the region of interest. Once these contiguous regions ("contigs") have been formed, the clones within them can then be sequenced. These sequences can then be searched for the presence of putative genic regions, and through a series of complex genetic studies involving mutations, knockouts, and transformations, the researcher may be lucky enough to determine the gene or genes affecting the phenotype controlled by the QTL.

There are numerous roadblocks to overcome if this method is to be used, particularly for organisms with complex genomes, such as maize, now thought to be an ancient tetraploid. The presence of highly repetitive, non-coding DNA sequences can make sequencing and sequence alignment difficult, if not impossible. In addition, many apparently genic sequences may actually be duplicate, non-functional genes or "pseudogenes" that function only as evolutionary artifacts. Also, many species of plants are resistant to transformation, making the confirmation of gene function using gene "knockouts" restrictive. One method for reducing the complexity of certain genomes is the use of syntenic relationships. If a closely related, but less complex organism exists that has been well-characterized genetically, then one could map BACs or YACs from the related species to the QTL in the species of interest. Sequencing of these low complexity, syntenic clones may reveal genes that could be putatively conserved enough that a homeologue could be found in the species of interest. Comparisons of QTLs in syntenic regions across species may confirm this hypothesis.

An additional means of identifying genes underlying QTLs involves the use of candidate genes (Faris et al., 1999). In order to utilize this approach, a well-educated "guess" must be made as to what biochemical pathway(s) might be involved in the trait affected by the QTL of interest. Knowledge of the sequences of at least some of the genes involved in the pathway, perhaps from studies in other organisms, would allow the researcher to determine, through mapping, if any of these genes are located in or near the QTL of interest. Mapping genes to QTLs only tags them as putative candidates. Extensive biochemical and genetic studies must follow to confirm the phenotypic effect. This might include the use of gene-knockout populations such as those created in maize using transposable elements (e.g., the TUSC system created by Pioneer HiBred International). The existence of a known transposon sequence within a gene will normally knock out the function of that gene. F_1 plants containing a transposon within a candidate gene can be identified by the polymerase chain reaction (PCR) using primers designed specifically for that gene and the appropriate transposon. By performing a segregation analysis on the F_2 seed from the identified plant, it should be possible to confirm the putative function of the gene.

The above comments should serve to emphasize that the identification of large chromosomal segments (QTLs) that have an effect on a trait is only the beginning of a long and arduous process to determine the underlying genes controlling the trait of interest. Marker-facilitated techniques provide valuable tools for the rapid transfer of known trait variability from one individual or population to another. However, to precisely manipulate a trait and/or to create variability that does not exist in natural populations, it is necessary to understand the structure and function of all genes involved in the expression of that trait. This requires a highly coordinated effort that includes large scale sequencing efforts, gene annotation, highly integrated genetic and physical mapping, and studies of syntenic relationships. This massive amount of information must then be carefully brought together into a comprehensive, interactive, accurate, highly accessible, and user-friendly database. This is when genomics becomes involved.

Physical and Genetic Maps

Genomics is a whole-genome approach where one focus is on developing dense physical and genetic maps.

It relies on high throughput technologies that have been advanced for human and microbial genome sequencing projects. In addition to highly efficient wet-laboratory methodologies, the massive quantities of data generated require robust databases, data mining, and analysis tools, which are collectively termed bioinformatics. The physical maps being constructed include (i) those consisting of complete sequences, anticipated for *Arabidopsis* by Year 2004 and planned for rice; (ii) clusters of overlapping BAC clones (contigs), proposed for maize, sorghum [*Sorghum bicolor* (L.) Moench], and rice; and (iii) larger cloned chromosomal segments, such as radiation hybrids. Radiation hybrid technology (Cox et al., 1990), where segments of chromosomes are isolated and maintained in a hamster (*Mesocricetus* spp.) or other cell line, has yet to be applied to plants. It is an inexpensive and powerful technology for generating a physical map, and for linking physical and genetic maps at much higher resolution than current genetic mapping procedures. Some questions exist, however, as to whether appropriate selectable markers can be developed and whether animal/plant protoplast fusions will be successful. Once created, physical maps are anchored to the genetic map by conventional marker technologies (as described in the first part of this paper), including newer marker types, such as simple sequence repeats (SSRs, microsatellites), along with radiation hybrids, and potentially other new technologies as they arise.

Sequencing Efforts

A key to uncovering all possible genes in an organism is whole genome sequencing. For many organisms, however, large genome sizes and highly repetitive DNA sequences may make this task prohibitive. One approach to sequencing genomes, but with some limitations, is to sequence randomly cDNAs or ESTs (Expressed Sequence Tags) from various tissue libraries. This provides rapid access to many of the gene sequences of an organism and is cost-effective, especially when compared with "targeted one-gene-at-a-time sequencing." Sequence analysis tools, using annotations from previous studies, provide many clues about functions of the sequences. However, it is quite possible that this approach will preferentially exclude any genes expressed in low copy number. If these are key regulatory loci that encode QTL variances, then clearly other strategies are required.

Completely sequencing a genome, even with the technological advances provided by the human genome project, remains relatively expensive. Much of the genome in many crop plants, including maize, is composed of non-coding regions with no apparent function. The strategic solution has been to sequence at least one model organism, *Arabidopsis*, and probably rice (National Academy of Sciences, 1998) to ensure that the sequences of all the genes of a higher plant are available. Researchers can then rely on dense genetic and physical maps for other plants, along with comparative mapping efforts to aid in discovering homologous genes in other species. The genes and their organization in quite distantly related species, such as wheat, maize and rice, are remarkably conserved (Ahn and Tanksley, 1993; Ahn et al., 1993; Kurata et al., 1994; Pereira and Lee, 1995; Van Deyne et al., 1995; Chen et al., 1997; Gale and Devos, 1998; Han et al., 1998). These comparative maps have been based largely on small sets of RFLP and isozyme data, although there is some evidence among mutants that lends support to genetic conservation (Doebley et al., 1997). With higher resolution maps, it is anticipated that the relationships will be defined even further, and exceptions to the patterns will be delineated. The goal of comparative genomics is to synthesize all of the information gained in a number of species and then use all of the available data to predict gene function across species. This will be critical to plants with large genomes, such as wheat and for evolutionarily related orphan species such as tef, *Oryza glaberrima* Steudel (African rice), and finger and pearl millets (*Pennisetum* spp.) which have smaller sets of map data.

Bioinformatics

Resources. Bioinformatics may be simply described as the data repositories, data mining, and analysis tools designed to interpret the genome data that is currently available along with the data that will soon deluge the plant science community. Current public repositories include sequence databases, species-specific genome databases and the germplasm databases. The sequence databases, such as GenBank (nucleotides) and Swiss-Prot (amino acids) are robust resources for sequence deposition and sequence analyses. They provide powerful on-line tools for sequence analysis and searching, where searches can be made for motifs and secondary structure as well as for amino acid or nucleotide similarities. Sequence databases can support record-to-record links to the species-specific databases. However, these links need to be specified by curators of the species-specific databases.

Species-specific databases exist for major crops, such as soybean, maize, wheat, rye (*Secale cereale* L.), barley, oat (*Avena sativa* L.), and rice, and for model organisms such as *Arabidopsis* and *Chlamydomonas*. These genome databases integrate the map data for the species, and provide documentation on the functionality of the genome. These data include the physical and genetic maps, clones and primers, QTLs, trait variances, references, images of pest and stress responses, and mutant phenotypes. They access the sequence information curated in the central sequence databases. The species-specific databases require scientific curation to ensure data quality, uniformity of gene and allele nomenclature, and accurate integration of data.

The Germplasm databases catalog information on available seed resources, along with certain agronomic and quality trait data. They presently do not contain any genome information, but linking to the genome databases is under investigation. GRIN, the U.S. Germplasm resource has initiated some links; the SINGER consortium, or System-wise Information Network for GEnetic Resources, is also considering the establish-

ment of such links. SINGER inputs data from major international germplasm banks in South America, Asia, Africa, the Middle East, and the Philippines. The Germplasm Enhancement of Maize (GEM) project is working with MaizeDB towards harmonizing the representation of value-added traits with reciprocal links to the evaluations of yield and various other traits. These traits are being transferred into U.S. corn belt germplasm from exotic lines.

Current Status. The maize genome may have nearly the complexity of the human genome. In maize, it is estimated that there are over 40 000 genes encoded in some 4.8 pg DNA. The same number of genes may also be encoded by rice (0.9 pg DNA) or by sorghum (1.6 pg DNA; Bennett et al., 1997). Currently, 4000 genes and probed sites (RFLP, SSR, RAPD) have map positions. There are 2000 designated genes, where fewer than half have been mapped to chromosomes or have been sequenced. Less than half of the total mapped genes and probed sites are ordered on any single map, with the BNL (1699) and UMC(1856) maps being most extensive. There are 390 loci shared between these two maps (Coe et al., 1998).

Query Example. A key to the effective utilization of the massive amount of information generated by the many genomics efforts will be the ease with which is can be accessed and analyzed. This must be accomplished through a series of logical data queries.

Consider that in some maize lines, the QTL with the largest effect on earworm resistance and the accumulation of maysin is associated with the *p1* pericarp color locus (Byrne et al., 1996). The *p1* locus is well characterized and has been sequenced. In designing lines with enhanced pest resistance, one might wish to ask whether there are other anthocyanin synthesis loci involved in insect response, either in maize or closely related species such as sorghum or rice. Because a comprehensive, cross-species data repository (GenBank) is available for sequences, one can readily, with a single query, discover all the sequences with similarity to the *p1* locus. With the species information in hand, in most instances, one could then query the species-specific databases or use the connecting link from GenBank if it has been established.

There are a number of genomics-based web sites that may need to be accessed for a single candidate gene identification study. A coordinated package which links together all of these resources would definitely facilitate this effort. In practice, such links are largely incomplete at this time. The following is a selected list of web sites that may be useful:

1. National Corn Growers Assoc. http://www.ncga.com/ (verified June 2, 1999).
2. NSF Plant Genome Project http://www.nsf.gov/ (verified June 2, 1999).
3. GenBank http://www.ncbi.nlm.nih.gov/ (verified June 2, 1999).
4. SwissProt http://expasy.hcuge.ch/ (verified June 2, 1999).
5. Plant Genome Databases http://probe.nalusda.gov/ (verified June 2, 1999).
6. MaizeDB http://www.agron.missouri.edu/ (verified June 2, 1999).
7. GRIN http://www.ars-grin.gov/ (verified June 2, 1999).
8. SINGER http://www.cgiar.org/iita/research/singer.htm (verified June 2, 1999).
9. GEM Project http://www.iastate.edu/~usda-gem/ (verified June 2, 1999).
10. Radiation Hybrids (Human) http://www.ebi.ac.uk/RHdb/index.html (verified June 2, 1999).
11. U.S. Rice Sequencing Project http://www.reeusda.gov/crgam/nri/programs/rice/rice.htm (verified June 2, 1999).

Integration of Marker Technology and Genomics into Plant Breeding Strategies

Reflections on Academic Investigations

Molecular-marker technology has been shown to be effective for identifying and mapping QTLs in numerous crop plants. Also, the positive results from a limited number of marker-facilitated selection and introgression studies are encouraging for transferring desired genes between breeding lines and, thus, increasing the precision and efficiency of plant breeding. Emerging technology should provide the vehicles for using markers to expedite the acquisition of important genes from exotic populations or from wild species.

In maize, for example, most mapping studies have been conducted in populations derived from domestic lines. Efforts in exotic maize populations have been less effective and frequently have met with considerable frustration (Koester, 1992). This has been particularly true in the use of RFLPs because of the large number of marker variants and multiple banding patterns that have been difficult to interpret. Differentiating between multiple alleles at a single marker locus versus alleles associated with duplicate loci frequently is nearly impossible. Hopefully, with new marker technology (e.g., PCR-based techniques), this limitation will be overcome and the use of exotic germplasm can be tapped as a vast source of new genes in maize breeding, as well as in other crops.

In most QTL identification studies, rather stringent probability levels have been set so that there is a low risk in making Type I errors (i.e., false positives). Thus, only QTLs with major effects are identified as being significant. It should be pointed out that these would be the QTLs with high heritabilities, are easily manipulated by traditional breeding practices, and may already be fixed in many breeding lines. It may prove to be more productive, therefore, to use marker technology as a means for placing greater emphasis on those QTLs (or chromosomal regions) that show only relatively minor effects.

Attempts have been made to reduce the size of the regions identified as containing major QTLs through "fine-mapping" in studies such as the one reported by Paterson et al. (1990). This approach has been envisioned as an initial step in identifying single genes that might ultimately be manipulated by transformation (re-

combinant DNA) technology. In fact, Alpert and Tanksley (1996) have used high resolution fine-mapping to isolate a YAC contig containing a major fruit weight QTL in tomato.

The identification and cloning of single genes could prove to be counter-productive, however, if these major QTLs contain a number of genes that have evolved as highly integrated (epistatic) complexes over many cycles of selection. For example, the region on the short arm of chromosome 5, bracketed by isozyme markers *Amp3* and *Pgm2*, has been found to have a very significant effect on grain yield in several maize populations (Stuber, 1992; Stuber et al., 1992). This region has been targeted for fine-mapping and it was found that there were at least two smaller QTLs (Graham et al., 1997). Although earlier analyses (Stuber et al., 1992) suggested that the major QTL identified in this region acted in an overdominant fashion, effects at these two smaller QTLs appear to act in a dominant manner, and are in repulsion phase linkage. Thus, the effects at these two QTLs support the dominance theory of heterosis in this chromosomal region. It will be of interest to determine whether chromosomal segments such as this can be further improved with marker technology (e.g., by placing the favorable dominant factors in coupling phase linkage in one, or both, of the parental lines). In practical breeding programs, manipulation of large segments (such as the one identified in maize on chromosome 5) may be simpler, and more effective in the short term, than extracting and manipulating individual genes.

It should be stressed that little is known regarding the stability of QTL alleles when transferred to different genetic backgrounds and when evaluated in varying environments. Tanksley and Hewitt (1988) illustrated the potential dangers of establishing breeding programs based on associations of markers with quantitative traits prior to evaluations of the identified factors in appropriate genetic backgrounds. Although there is some evidence for the interaction of QTLs with environments, results tend to be contradictory. Stuber et al. (1992) found little evidence for such interaction in maize. In a major follow-up study in a population generated from the cross of maize inbred lines B73 and Mo17 (Stuber et al., 1996; LeDeaux and Stuber, 1997; Stuber, 1999, 1998; and 1996, unpublished data), there was little evidence for QTL $\times$ environment interaction even when very severe stress conditions were imposed on the field evaluations. These data would suggest that breeders and geneticists can rely on mapping data from favorable environments for breeding materials adapted to stress environments. This suggestion should be viewed with caution, however, because the results may reflect the fact that the parental lines have been selected for stability over a wide range of environments, and it may not be prudent to extrapolate to more widely variable and more divergent materials.

Empirical Breeding Strategies

As Tanksley and Nelson (1996) point out, the impact of marker-based QTL analysis on the development of new lines or varieties with enhanced quantitative traits has been less than many have expected. They list two primary reasons: (i) the discovery of QTLs and variety development have been two separate processes, and (ii) most breeding-related QTL studies have been targeted toward the manipulation of quantitative traits in elite germplasm. We would add another: (iii) for traits such as grain yield, QTL expression usually is dependent upon the genetic background in which it is found; therefore QTL evaluation must be done independently each time a new population or cross is used. For less complex traits, such as disease or insect resistance, this is usually not the case, however.

About 10 yr ago, in our maize research program at Raleigh, we designed a study using an alternative approach in which line development and QTL identification were combined into a single integrated process. This is discussed earlier in this paper in the section: "Near-Isogenic Lines (NILs) as a Breeding Tool." The procedure requires no advance QTL discovery. In fact, QTL identification is a fringe benefit from the breeding process. It is useful for introducing new alleles from other elite breeding lines or from exotic or unadapted germplasm sources. We have used it successfully to introduce favorable alleles from an exotic maize population into the elite inbred line, Mo17 (Furbeck, 1993).

A somewhat similar integrated approach, advanced backcross-QTL analysis (AB-QTL) has been proposed by Tanksley and Nelson (1996). It differs from our NIL approach in that the AB-QTL scheme uses traditional backcrossing through the BC_2 or BC_3 generation, with phenotypic selection to eliminate visually deleterious factors. A goal of the NIL approach is to develop a series of NILs, each of which contains a different chromosomal segment from the donor parent, but collectively encompasses all of the genetic material from the donor. Because the AB-QTL scheme uses traditional backcrossing, the resulting lines depend on a stochastic approach, and the probability of collectively encompassing all of the genetic material from the donor (i.e., having each of the donor chromosomal segments represented in at least one of the BC lines) is very low (see Beckmann and Soller, 1986) unless several thousand lines are generated and evaluated. In addition, if the goal is to improve a trait such as yield, phenotypic selection could be counter-productive. Favorable yield genes may be linked to the genes associated with the selected traits, and would be eliminated during the backcrossing process. If this approach had been used in the study by Furbeck (1993), a favorable yield gene (contributing 659 kg ha^{-1}; 10.5 bushels/acre) closely linked to a gene associated with a deleterious trait would probably have been undetected.

It should be pointed out that modifications to the NIL scheme can be made to reduce the costs involved in the marker analyses. For example, the breeding material could be grouped into subsets; each subset would then be involved with marker analyses for only one, or possibly two, chromosomes. In maize with 10 chromosomes, this might involve 10 subsets. Thus, in each subset, each of the NILs developed would contain a different donor

segment for that specific chromosome. By increasing the number of backcrosses to five or six, the recipient genome should be essentially recovered in the other nine chromosomes. With three generations in the breeding nurseries per year, this development process could be accomplished in 2 yr.

Obviously, the goals for any of these breeding strategies are to identify and introgress useful genetic factors into targeted breeding materials more effectively than can be accomplished by traditional breeding approaches. Genomics will eventually be used to fine tune strategies such as those outlined above. However, as Paterson (1998b) stated, "manipulation of the vast majority of QTLs in plant breeding programs does not require cloning—so why bother?"

Conclusions

New investigations using DNA-based marker technology as a tool for plant geneticists and plant breeders continue to add evidence to the projected role of markers not only for identifying useful genes (or chromosomal segments) in various germplasm sources but also for transferring these genes into desired cultivars or lines. As laboratory analyses become more automated, the cost (now one of the major deterrents in the use of marker technology) will decrease and the use of DNA-based markers for improving multigenic traits will be greatly expanded in the future.

When genomics is added to future strategies for plant and animal breeders, the projected outcomes are mind-boggling. There is every reason to believe that the synergy of empirical breeding, marker-assisted selection, and genomics will truly "produce a greater effect than the sum of the various individual actions."

REFERENCES

Abler, B.S.B., M.D. Edwards, and C.W. Stuber. 1991. Isoenzymatic identification of quantitative trait loci in crosses of elite maize hybrids. Crop Sci. 31:267–274.

Ahn, S., J.A. Anderson, M.E. Sorrells, and S.D. Tanksley. 1993. Homeologous relationships of rice, wheat and maize chromosomes. Mol.Gen.Genet. 241:483–490.

Ahn, S., and S. Tanksley. 1993. Comparative linkage maps of the rice and maize genomes. Proc. Natl. Acad. Sci. (USA) 90:7980–7984.

Alpert, K.B., and S.D. Tanksley. 1996. High-resolution mapping and isolation of a yeast artificial chromosome contig containing fw2.2: A major fruit weight quantitative trait locus in tomato. Proc. Natl. Acad. Sci. (USA) 93:15503–15507.

Beckmann, J.S., and M. Soller. 1986. Restriction fragment length polymorphisms in plant genetic improvement. Oxford Surv. Plant Mol. Cell Biol. 3:197–250.

Bennett, M.D., A.V. Cox, and Leitch. 1997. Angiosperm DNA C-values database. http://www.rbgkew.org.uk/cval/ (verified June 2, 1999).

Bernardo, R. 1994. Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Sci. 34:20–25.

Brown, A.H.D., J. Munday, and R.N. Oram. 1989. Use of isozyme-marked segments from wild barley (*Hordeum spontaneum*) in barley breeding. Plant Breed. 100:280–288.

Bubeck, D.M., M.M. Goodman, W.D. Beavis, and D. Grant. 1993. Quantitative trait loci controlling resistance to gray leaf spot in maize. Crop Sci. 33:838–847.

Burow, M.D., and T.K. Blake. 1998. Molecular tools for the study of complex traits. p. 13–29. In A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton, FL.

Byrne, P.F., M.D. McMullen, M.E. Snook, T.A. Musket, J.M. Theuri, N.W. Widstrom, B.R. Wiseman, and E.H. Coe. 1996. Quantitative trait loci and metabolic pathways: Genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. Proc. Natl. Acad. Sci. (USA) 93:8820–8825.

Chen, M., P. SanMiguel, A.C. de Oliviera, S.-S. Woo, H. Shang, R.A. Wing, and J.L. Bennetzen. 1997. Microcolinearity in she-homologous regions of the maize, rice and sorghum genomes. Proc. Natl. Acad. Sci. (USA) 94:3431–3435.

Coe, E.H., M.L. Polacco, M.D. McMullen, and G.D. Davis. 1999. Maize molecular maps: Markers, bins and databases. *In* R. Phillips and I. Vasil (ed.) DNA-based markers in plants. Kluwer Academic Publishers, Dordrecht, the Netherlands (in press).

Cox, D.R., M.N. Burmeister, E.R. Price, S. Kim, and R.M. Myers. 1990. Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosome. Science 250:245–250.

Comstock, R. 1978. Quantitative genetics in maize breeding. p. 191–206. *In* D. Walden (ed.) Maize breeding and genetics. Wiley, New York.

Darrah, L.L., and A.R. Hallauer. 1972. Genetic effects estimated from generation means in four diallel sets of maize inbreds. Crop Sci. 12:615–621.

Doebley, J., A. Stec., and L. Hubbard. 1997. The evolution of apical dominance in maize. Nature 386:485–488.

Dudley, J.W., M.A. Saghai Maroof, and G.K. Rufener. 1992. Molecular marker information and selection of parents in corn breeding programs. Crop Sci. 32:301–304.

Duvick, D.N. 1997. What is yield? p. 332–335. *In* G.O. Edmeades et al. (ed.) Proc. of Symposium on Developing Drought- and Low N-Tolerant Maize. CIMMYT, El Batan, Mexico.

Eathington, S.R., J.W. Dudley, and G.K. Rufener. 1997. Marker effects estimated from testcrosses of early and late generations of inbreeding in maize. Crop Sci. 37:1679–1685.

Edwards, M.D., T. Helentjaris, S. Wright, and C.W. Stuber. 1992. Molecular-marker-facilitated investigations of quantitative trait loci in maize: IV. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. Theor. Appl. Genet. 83:765–774.

Edwards, M., and L. Johnson. 1994. RFLPs for rapid recurrent selection. p. 33–40. *In* Proc. of Symposium on Analysis of Molecular Marker Data. Am. Soc. Hort. Sci. and CSSA, Corvallis, OR.

Edwards, M.D., C.W. Stuber, and J.F. Wendel. 1987. Molecular marker facilitated investigations of quantitative trait loci in maize: I. Numbers, genomic distribution and types of gene action. Genetics 116:113–125.

Eshed, Y., and D. Zamir. 1995. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL. Genetics 141: 1147–1162.

Faris, J.D., W.L. Li, D.J. Liu, P.D. Chen, and B.S. Gill. 1999. Candidate gene analysis of quantitative disease resistance in wheat. Theor. Appl. Genet. 98:219–225.

Furbeck, S.M. 1993. The development and evaluation of molecular-marker derived near isogenic lines to study quantitative traits in maize. Ph.D. thesis, North Carolina State University, Raleigh (Diss. Abstr. DA9330287).

Gale, M.D., and K.M. Devos. 1998. Plant comparative genetics after 10 years. Science 282:656–659.

Gimelfarb, A., and R. Lande. 1994a. Simulation of marker-assisted selection in hybrid populations. Genet. Res. 63:39–47.

Gimelfarb, A., and R. Lande. 1994b. Simulation of marker-assisted selection for non-additive traits. Genet. Res. 64:127–136.

Gimelfarb, A., and R. Lande. 1995. Marker-assisted selection and marker-QTL associations in hybrid populations. Theor. Appl. Genet. 91:522–528.

Graham, G.I., D.W. Wolff, and C.W. Stuber. 1997. Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping. Crop Sci. 37:1601–1610.

Han, F., A. Kleinhofs, S.E. Ullrich, A. Kilian, M. Yano, and T. Sasaki. 1998. Synteny with rice: analysis of barley malting quality QTLs and rpg4 chromosome regions. Genome 41:373–380.

Hospital, F., L. Moreau, F. Lacoudre, A. Charcosset, and A. Gallais. 1997. More on the efficiency of marker-assisted selection. Theor. Appl. Genet. 95:1181–1189.

Johnson, G.R., and R.H. Mumm. 1996. Marker assisted maize breeding. Proc. 51st Annual Corn and Sorghum Industry Research Conf., American Seed Trade Assoc. 51:75–84.

Kearsey, M.J., and A.G.L. Farquhar. 1998. QTL analysis in plants; where are we now? Heredity 80:137–142.

Knapp, S.J. 1994. Mapping quantitative trait loci. p. 58–96. *In* R. Phillips and I. Vasil (ed.) DNA-based markers in plants. Kluwer Academic Publishers, Dordrecht, the Netherlands.

Knapp, S.J. 1998. Marker-assisted selection as a strategy for increasing the probability of selecting superior genotypes. Crop Sci. 38: 1164–1174.

Koester, R.P. 1992. Identification of quantitative trait loci controlling maturity and plant height using molecular markers in maize (*Zea mays* L.) Ph.D. thesis (Diss. Abstr. AAI9221518). North Carolina State University, Raleigh.

Koester, R.P., P.H. Sisco, and C.W. Stuber. 1993. Identification of quantitative trait loci controlling days to flowering and plant height in two near isogenic lines of maize. Crop Sci. 33:1209–1216.

Kurata, N., G. Moore, Y. Nagamura, T. Foote, M. Yana, Y. Minobe, and M.D. Gale. 1994. Conservation of genome structure between rice and wheat. Bio/Technology 12:276–278.

National Academy of Sciences. 1998. June 1997, Irvine CA, Colloquium, Protecting our food supply: The value of plant genome initiatives. Proc. Natl. Acad. Sci. (USA) 95:1969–2032.

Lande, R., and R. Thompson. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756.

LeDeaux, J.R., and C.W. Stuber. 1997. Mapping heterosis QTLs in maize grown under various stress conditions. Symposium on the Genetics and Exploitation of Heterosis in Crops. Abstracts. p. 40–41.

Li, Z., S.R.M. Pinson, W.D. Park, A.H. Paterson, and J.W. Stansel. 1997. Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics 145:453–465.

Melchinger, A.E., M. Lee, K.R. Lamkey, and W.L. Woodman. 1990. Genetic diversity for restriction fragment length polymorphisms: Relation to estimated effects in maize inbreds. Crop Sci. 30: 1033–1040.

Paterson, A.H. 1998a. High resolution mapping of QTLs. p. 163–173. *In* A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton, FL.

Paterson, A.H. 1998b. Prospects for cloning the genetic determinants of QTLs. p. 289–293. *In* A.H. Paterson (ed.) Molecular dissection of complex traits. CRC Press, Boca Raton, FL.

Paterson, A.H., J.W. Deverna, B. Lanini, and S.D. Tanksley. 1990. Fine mapping of quantitative trait loci using selected overlapping recombinant chromosomes, in an interspecies cross of tomato. Genetics 124:735–742.

Paterson, A.H., S. Damon, J.D. Hewitt, D. Zamir, H.D. Rabinowitch, S.E. Lincoln, E.S. Lander, and S.D. Tanksley. 1991. Mendelian factors underlying quantitative traits in tomato. Comparison across species, generations, and environments. Genetics 127:181–197.

Peng, J.Y., J.C. Glaszmann, and S.S. Virmani. 1988. Heterosis and isozyme divergence in indica rice. Crop Sci. 28:561–563.

Pereira, M.G., and M. Lee. 1995. Identification of genomic regions affecting plant height in sorghum and maize. Theor. Appl. Genet. 90:380–388.

Ragot, M., P.H. Sisco, D.A. Hoisington, and C.W. Stuber. 1995. Molecular-marker-mediated characterization of favorable exotic alleles at quantitative trait loci in maize. Crop Sci. 35:1306–1315

Rector, B.G., J.N. All, W.A. Parrott, and H.R. Boerma. 1998. Identification of molecular markers linked to quantitative trait loci for soybean resistance to corn earworm. Theor. Appl. Genet. 96: 786–790.

Rector, B.G., J.N. All, W.A. Parrott , and H.R. Boerma. 1999. Quantitative trait loci for antixenosis to corn earworm in soybean. Crop Sci. 39:531–538.

Smith, J.S.C., and O.S. Smith. 1989. The use of morphological, biochemical, and genetic characteristics to measure distance and to test for minimum distance between inbred lines of maize (*Zea mays* L.) (Mimeo of paper presented at UPOV Workshop, Versailles, France, October 1989). Pioneer Hi-Bred International, Inc., Johnston, IA.

Stuber, C.W. 1992. Biochemical and molecular markers in plant breeding. *In* J. Janick (ed.) Plant Breeding Reviews 9:37–61.

Stuber, C.W. 1994. Success in the use of molecular markers for yield enhancement in corn. Proc. 49th Annual Corn and Sorghum Industry Res. Conf., American Seed Trade Assoc. 49:232–238.

Stuber, C.W. 1995. Mapping and manipulating quantitative traits in maize. Trends in Genetics 11:477–481.

Stuber, C.W. 1999. Breeding multigenic traits. *In* R. Phillips and I. Vasil (ed.) DNA-based markers in plants. Kluwer Academic Publishers, Dordrecht, the Netherlands (in press).

Stuber, C.W. 1998. Case history in crop improvement: Yield heterosis in maize. p. 197–206. *In* A.H. Paterson (ed.), Molecular Analysis of Complex Traits. CRC Press, Boca Raton, FL.

Stuber, C.W., and M.D. Edwards. 1986. Genotypic selection for improvement of quantitative traits in corn using molecular marker loci. Proc. 41st Annual Corn and Sorghum Industry Res. Conf., American Seed Trade Assoc. 41:70–83.

Stuber, C.W., M.D. Edwards, and J.F. Wendel. 1987. Molecular marker facilitated investigations of quantitative trait loci in maize: II. Factors influencing yield and its component traits. Crop Sci. 27:639–648.

Stuber, C.W., G. Graham, and M.L. Senior. 1996. Effects of environmental stresses on mapping heterosis QTLs in maize. Plant Genome IV Abstr. p. 23.

Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. Genetics 132:823–839.

Stuber, C.W., and P.H. Sisco. 1991. Marker-facilitated transfer of QTL alleles between elite inbred lines and responses in hybrids. Proc. 46th Annual Corn and Sorghum Industry Res. Conf., American Seed Trade Assoc. 46:104–113.

Tanksley, S.D., and J. Hewitt. 1988. Use of molecular markers in breeding for soluble solids content in tomato—a re-examination. Theor. Appl. Genet. 75:811–823.

Tanksley, S.D., and J.C. Nelson. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. 92:191–203.

Van Deyne, A.E., J.C. Nelson, S.E. Harrington, E.S. Yglesias, D. Braga, S.R. McCouch, and M.E. Sorrells. 1995. Comparative mapping in grasses. Mol. Gen. Genet. 248:744–754.

Wittaker, J.C., R.N. Curnow, C.S. Haley, and R. Thompson. 1995. Using marker-maps in marker-assisted selection. Genet. Res. 66: 255–265.

Yamamoto, T., Y. Kuboki, S.Y. Lin, T. Sasaki, and M. Yano. 1998. Fine mapping of quantitative trait loci *Hd-1*, *Hd-2* and *Hd-3*, controlling heading date of rice, as single Mendelian factors. Theor. Appl. Genet. 97:37–44.

Zhang, W., and C. Smith. 1992. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. Theor. Appl. Genet. 83:813–820.

Zhang, W., and C. Smith. 1993. Simulation of marker-assisted selection utilizing linkage disequilibrium: the effects of several additional factors. Theor. Appl. Genet. 86:492–496.